

VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ

Fakulta elektrotechniky
a komunikačních technologií

DIPLOMOVÁ PRÁCE



VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ

BRNO UNIVERSITY OF TECHNOLOGY

**FAKULTA ELEKTROTECHNIKY
A KOMUNIKAČNÍCH TECHNOLOGIÍ**

FACULTY OF ELECTRICAL ENGINEERING AND COMMUNICATION

ÚSTAV TELEKOMUNIKACÍ

DEPARTMENT OF TELECOMMUNICATIONS

**ANALÝZA PŮVODU VZNIKU INTERPRETACE ZA POMOCI
METOD PARAMETRIZACE HUDEBNÍHO SIGNÁLU**

DETERMINING THE PLACE OF ORIGIN OF INTERPRETATION BASED ON THE MUSIC SIGNAL
PARAMETERIZATION

DIPLOMOVÁ PRÁCE

MASTER'S THESIS

AUTOR PRÁCE

AUTHOR

Bc. Tomáš Brada

VEDOUCÍ PRÁCE

SUPERVISOR

Ing. Tomáš Kiska

BRNO 2019

Diplomová práce

magisterský navazující studijní obor **Audio inženýrství**

Ústav telekomunikací

Student: Bc. Tomáš Brada

ID: 154683

Ročník: 2

Akademický rok: 2018/19

NÁZEV TÉMATU:

Analýza původu vzniku interpretace za pomoci metod parametrizace hudebního signálu

POKYNY PRO VYPRACOVÁNÍ:

Shrňte dosavadní poznatky z oblasti srovnávání zvukových nahrávek v oblasti tzv. „Music information retrieval“. Dále analyzujte interpretace z hlediska tempa, dynamiky a barvy zvuku. Implementujte systém, který za pomoci metod strojového učení a korelační analýzy vybere takové hudební parametry, které mají největší schopnost diferencovat původ vzniku jednotlivých interpretací.

DOPORUČENÁ LITERATURA:

[1] SCHEDL, M.; GOMÉZ, E; URBANO, J. Music Information Retrieval: Recent Developments and Applications. Foundations and Trends in Information Retrieval, 2014: s. 127-261, DOI: 978-1-60198-807-2.

[2] MÜLLER, M. Fundamentals of Music Processing: Audio, Analysis, Algorithms, Applications [online]. Springer International Publishing Switzerland, 2015, 483 s. ISBN 978-3-319-21945-5.

Termín zadání: 1.2.2019

Termín odevzdání: 16.5.2019

Vedoucí práce: Ing. Tomáš Kiska

Konzultant:

prof. Ing. Jiří Mišurec, CSc.
předseda oborové rady

UPOZORNĚNÍ:

Autor diplomové práce nesmí při vytváření diplomové práce porušit autorská práva třetích osob, zejména nesmí zasahovat nedovoleným způsobem do cizích autorských práv osobnostních a musí si být plně vědom následků porušení ustanovení § 11 a následujících autorského zákona č. 121/2000 Sb., včetně možných trestněprávních důsledků vyplývajících z ustanovení části druhé, hlavy VI. díl 4 Trestního zákoníku č.40/2009 Sb.

ABSTRAKT

Tato práce shrnuje dosavadní poznatky z oblasti srovnávání zvukových nahrávek v oblasti tzv. „Music information retrieval“. Dále je cílem této práce analyzovat interpretace z hlediska tempa, dynamiky a barvy zvuku. Ze získaných dat provedené analýzy je pak cílem vybrat takové hudební parametry, které mají největší schopnost diferencovat původ vzniku jednotlivých interpretací.

KLÍČOVÁ SLOVA

Tempo, barva zvuku, dynamika, hudební parametry, získávání hudební informace, MIR-toolbox, Matlab, interpretace

ABSTRACT

Presented thesis summarizes existing findings in the problematics of music recording comparison, in the field of so-called Music Information Retrieval. One of the goals of the thesis is to perform analysis of multiple musical renditions from the point of view of the tempo, dynamics and sound timbre. From the obtained data, such musical parameters will be chosen which exhibit the biggest potential to differentiate the origine of the individual renditions.

KEYWORDS

Tempo, timbre, dynamics, musical features, music information retrieval, MIRtoolbox, Matlab, interpretation

BRADA, Tomáš. *Analýza původu vzniku interpretace za pomoci metod parametrizace hudebního signálu*. Brno, Rok, 57 s. Diplomová práce. Vysoké učení technické v Brně, Fakulta elektrotechniky a komunikačních technologií, Ústav telekomunikací. Vedoucí práce: Ing. Tomáš Kiska,

PROHLÁŠENÍ

Prohlašuji, že svou diplomovou práci na téma „Analýza původu vzniku interpretace za pomoci metod parametrizace hudebního signálu“ jsem vypracoval(a) samostatně pod vedením vedoucího diplomové práce a s použitím odborné literatury a dalších informačních zdrojů, které jsou všechny citovány v práci a uvedeny v seznamu literatury na konci práce.

Jako autor(ka) uvedené diplomové práce dále prohlašuji, že v souvislosti s vytvořením této diplomové práce jsem neporušil(a) autorská práva třetích osob, zejména jsem nezasáhl(a) nedovoleným způsobem do cizích autorských práv osobnostních a/nebo majetkových a jsem si plně vědom(a) následků porušení ustanovení § 11 a následujících autorského zákona č. 121/2000 Sb., o právu autorském, o právech souvisejících s právem autorským a o změně některých zákonů (autorský zákon), ve znění pozdějších předpisů, včetně možných trestněprávních důsledků vyplývajících z ustanovení části druhé, hlavy VI. díl 4 Trestního zákoníku č. 40/2009 Sb.

Brno

.....

podpis autora(-ky)

PODĚKOVÁNÍ

Rád bych poděkoval vedoucímu diplomové práce panu Ing. Tomášovi Kiskovi za odborné vedení, konzultace, trpělivost a podnětné návrhy k práci.

Brno

.....

podpis autora(-ky)

PODĚKOVÁNÍ

Výzkum popsany v této diplomové práci byl realizován v laboratořích podpořených z projektu SIX; registrační číslo CZ.1.05/2.1.00/03.0072, operační program Výzkum a vývoj pro inovace.

Brno

.....
podpis autora(-ky)

OBSAH

Úvod	11
1 Základní hudební a zvukové vlastnosti	12
1.1 Tempo	12
1.2 Dynamika	14
1.3 Barva zvuku	15
1.4 Artikulace	15
2 Rozbor skladby	18
2.1 Impresionismus	18
2.2 Claude Debussy	18
2.2.1 Život a dílo	18
2.2.2 Interpretace a technika klavírní hry	19
2.3 Svit luny	20
3 Získávání hudební informace	22
3.1 Parametrizace	22
3.1.1 Rytmické a tempové parametry	23
3.1.2 Dynamické parametry	25
3.1.3 Parametry barvy zvuku	25
3.1.4 Statistický popis spektrálního rozložení	28
4 Klasifikace	30
4.1 Rozhodovací strom	30
4.2 Křížová validace	31
4.3 Minimální redundance, maximální relevance	31
4.4 Vyhodnocování	32
4.4.1 Přesnost	32
4.4.2 Senzitivita	33
4.4.3 Specifita	33
4.4.4 Kompromis mezi senzitivitou a specifitou	33
5 Návrh analýzy nahrávek	34
5.1 Databáze	34
5.2 Předzpracování	34
5.3 MIRtoolbox a extrakce parametrů	35
5.4 Vyhodnocování a porovnání	35

6 Závěr	39
Literatura	40
Seznam symbolů, veličin a zkratk	46
Seznam příloh	47
A Tabulka klavírních interpretů	48
B Notový zápis	49
C Tabulka parametrů	55
D Obsah přiloženého CD	57

SEZNAM OBRÁZKŮ

1.1	Značení tempa v notovém zápise	13
1.2	Značení dynamiky v notovém zápise	14
1.3	Řada alikvotních tónů	15
1.4	Spektrum tónu A1 zahrané klavírem	16
1.5	Značení artikulace v notovém zápise	17
3.1	Matice podobnosti	24
3.2	ADSR obálka	26
4.1	Rozhodovací binární strom	31
5.1	Návrh analýzy nahrávek	34

SEZNAM TABULEK

1.1	Názvosloví temp	12
1.2	Označení dynamiky	14
5.1	Tabulka výsledků	36
A.1	Tabulka interpretů	48
C.1	Tabulka parametrů 1. část	55
C.2	Tabulka parametrů 2. část	56

ÚVOD

Hudba doprovází člověka od pravěku až po současnost. A právě v současné informační době s vývojem výpočetní techniky, která nám umožňuje odkrývat informace skryté ve zvukových nahrávkách, které jsou člověkem vědomě nepostřehnutelné.

Tato práce se zabývá popisem hudebních vlastností jako je tempo, dynamika a barva zvuku jak z pozice hudebníka, který čte informace z notového zápisu, tak i z pozice technika, který analyzuje digitální zvukový signál. Dále se věnuje vývoji strojového učení a umělé inteligenci, která dokáže vnímat hudbu podobně jako člověk. Dokáže ze základních vlastností a parametrů získaných z hudební nahrávky identifikovat použité akordy, určit tóninu, rozpoznat instrumentační obsazení, odhadnout původ a období vzniku skladby a zařadit ji do příslušného uměleckého směru. Dokáže také rozdělit hudbu podle žánru, nálady nebo sociálního kontextu. Tomuto odvětví bádání se ujal název „Music information retrieval“.

Výstupem této práce je návrh analýzy hudebních nahrávek, který by umožnil pomocí získaných parametrů určit původ interpretace impresionistické skladby *Clair de lune*.

1 ZÁKLADNÍ HUDEBNÍ VLASTNOSTI

1.1 Tempo

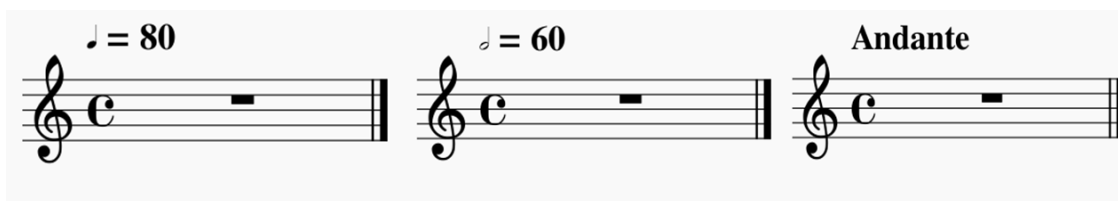
Tempo (*tempus* – čas) je významným parametrem, který zachycuje závislost posloupnosti not v čase. Jinými slovy jak rychle se má daná skladba zahrát. Pro vyjádření tempa v hudbě existují dva způsoby. Prvním z nich je označení pomocí noty (např. čtvrtová, půlová, ...) a číselné hodnoty na začátku notového zápisu. Určení tempa pomocí čísla je jednoznačné. Číslo určuje kolikrát se délka dané noty vejde do jedné minuty. Jednotkou je BPM (Beat Per Minute) – počet úderů za minutu. Druhým způsobem je určení tempa slovem nebo souslovím. V tomto případě není určení tempa zcela jednoznačné. Stejné označení může u různých skladatelů odpovídat odlišným úmyslům. Taktéž rychlost tempa záleží na interpretovi. Musí se rozhodnout na základě svých vědomostí o žánru, stylu, historického zasazení a svého hudebního cítění. Interpretace tempa skladby se tedy může lišit. Slovní popis tempa se obdobně jako v předchozím případě nachází na začátku notového zápisu. Pro označení tempa slovem se používá italské názvosloví (jen výjimečně se označuje v angličtině). V tabulce 1.1 jsou běžně se vyskytující názvosloví i s přibližnými hodnotami BPM. Pro postupné zrychlování se pak používá slovo *accelerando* (zkr. *accel.*), pro zpo-

Tab. 1.1: Názvosloví temp

Název	Překlad	BPM
grave	velmi pomalu	25 až 45
largo	široce	40 až 60
larghetto	zdlouhavě, ale rychleji než largo	60 až 66
adagio	pomalu	66 až 76
andante	zvolna, krokem	76 až 108
moderato	středně	108 až 120
allegro	rychle	120 až 168
vivace	živě	168 až 176
presto	velmi rychle	168 až 200
prestissimo	co nejrychleji	více než 200

malování *ritardando* (zkr. *rit.*), *ritenuto* (zkr. *riten.*) nebo *rallentando* (zkr. *rall.*). Tempo I (*tempo primo*) je pak označení v místě skladby, kdy se má interpret vrátit k původnímu tempu po předešlém zrychlení či zpomalení.[45, 12, 32]

Tempo rubato (*ukradený čas*) je způsob hry, kdy interpret zrychluje nebo zpomaluje tempo podle jeho vlastního hudebního cítění. Existují dva druhy tempa rubata.



Obr. 1.1: Značení tempa v notovém zápise

Jeden je zcela volný a druhý je omezen tím, že celkové tempo jednoho taktu zůstává stejné, avšak uvnitř taktu se tempo mění. Jinými slovy tempo se v taktu mění tak, že některé tóny se zahrají rychleji a některé pomaleji – celkové tempo taktu zůstává stejné. Jedna nota „okrádá“ druhou notu o čas.[15]

Milan Kuna a Miloš Bláha ve své studii Čas a Hudba [21] rozdělují tempo do tří skupin a to:

1. tempo s přísnou vazbou hudebního tempa, tzv. tempo metrizované – předpokládá sled metrických struktur, v nichž se střídají teze a arze, přízvukné a nepřízvukné doby taktu.
2. Tempo s libovольnou vazbou hudebního metra – tempo není předepsáno, chybí i takty. Záleží jen na interpretovi, jakým způsobem bude jednotlivé časové úseky artikulovat, jaké metro jim přisoudí. Stále se však podléhá rytmizovanému pohybu hudby.
3. Tempo bez vazby na hudební metrum, tzv. tempo nemetrizované – tempo se osvobozuje od hudebního metra, soustřeďuje se na vteřinové odplývání času. Notový zápis je nahrazen grafickým znázorněním a časovým údajem ve vteřinách.

Tempo z každé skupiny se dá dále rozdělit na:

1. Základní tempo – je nadřazeno všem ostatním typům tempa. Vyjadřuje průměrnou rychlost tempa hudby bez ohledu na dílčí agogické odchylky, které hudebník uplatňuje ve své hře.
2. Pevné tempo – je charakterizováno časovým pohybem hudby, v němž nedochází k žádným agogickým výkyvům, ať už úmyslným či neúmyslným.
3. Oscilující tempo – oscilace kolem průměrné doby taktu. Většinou jsou to nahodilé tempové odchylky hrajícího hudebníka, která je způsobena technickou nedokonalostí hry nebo jeho agogickým cítěním.
4. Proměnné tempo – je tempo se záměrnými odchylkami (*tempo rubato*, *ritardando*, *accelerando*, *koruna nad notou*).

1.2 Dynamika

V tabulce 1.2 se nacházejí značky, které se používají v notovém zápise a v hudební teorii k popisu dynamiky. Pro postupné zesilování nebo zeslabování se používají

Tab. 1.2: Označení dynamiky

Označení	Název	Překlad
ppp	piano pianissimo	co nejslaběji
pp	pianissimo	velmi slabě
p	piano	slabě
mp	mezzopiano	středně slabě
mf	mezzoforte	středně silně
f	forte	silně
ff	fortissimo	velmi silně
fff	forte fortissimo	co nejsilněji

buďto znaky $<>$ nebo slova *crescendo* (zkr. *cresc.*) pro zesilování a *decrescendo* (zkr. *decresc.*) nebo *diminuendo* (zkr. *dim.*) pro zeslabování. Označení **sf** – *sforzato* nebo také **sfz** – *sforzando* předepisují velmi silný a krátký důraz jednoho tónu.[45] Tato



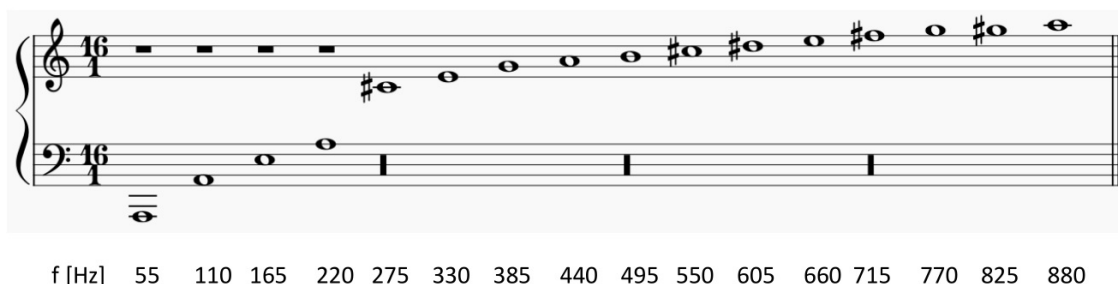
Obr. 1.2: Značení dynamiky v notovém zápise

hudební označení však určují dynamiku skladby jen orientačně. Skutečná hlasitost skladby je ovlivněna dalšími faktory např. instrumentací (rozdíl hlasitosti *forte* hrané příčnou flétnou nebo trumpetou), počtem hlasů (*forte* hrané klavírem nebo symfonickým orchestrem) či interpretací hráče.

V audiotechnice se pro indikaci úrovně signálu používá několik standardů. Nejpoužívanější z nich je Standard Volume Indicator (SVI) nebo Volume Unit (VU). Pro zobrazení úrovně signálu se používají VU metry nebo se spočítá a zobrazí amplitudová obálka.[44] Viz kapitola 3.1.2

1.3 Barva zvuku

Každý tón se skládá z tzv. fundamentu a vyšších harmonických. Frekvence fundamentu f udává výšku tónu. Vyšší harmonické jsou celočíselné násobky fundamentu tedy $2f$, $3f$, $4f$, atd. Dvojnásobek frekvence fundamentu nazýváme druhou harmonickou. Podle toho jakou mírou jsou vyšší harmonické v tónu obsaženy, určují barvu zvuku. Díky tomuto jsme schopni rozlišit zvuk houslí od kytary. Barvu zvuku nástroje lze měnit i technikou hry. Kytara vyzařuje různé barvy tónu, jestliže hráč hraje prsty, nehty či trsátkem. Taktéž lze barvu nástroje ovlivnit intenzitou a místem drnknutí na struně. Pro zobrazení frekvenčního spektra tónu se ve výpočetní



Obr. 1.3: Řada alikvotních tónů

technice používá diskretní Fourierova transformace (DFT) dána vztahem

$$F[k] = \sum_{n=0}^{N-1} f[n] e^{-j k \frac{2\pi}{N} n}, \quad k = 0, 1, \dots, N-1 \quad (1.1)$$

a pro inverzní diskretní Fourierovu transformaci (IDFT) vztahem

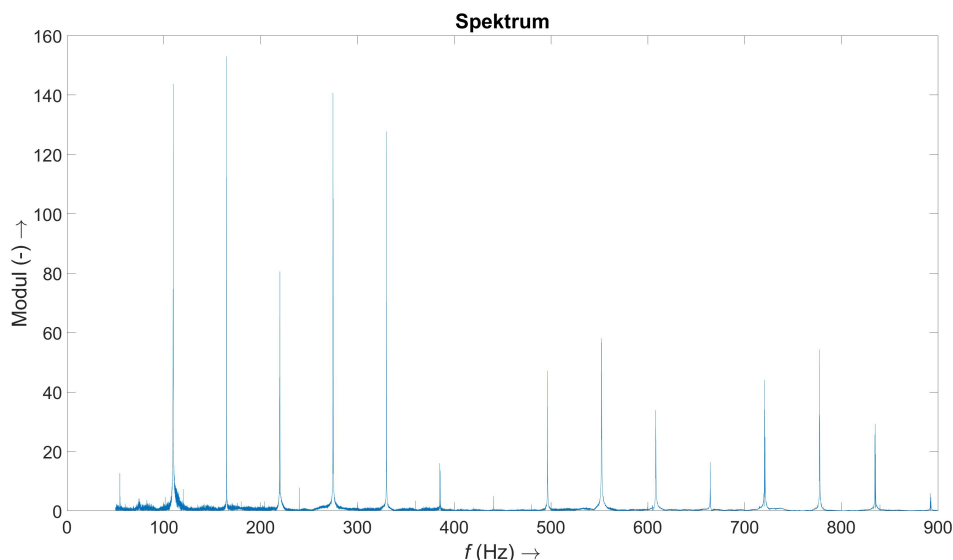
$$f[n] = \frac{1}{N} \sum_{k=0}^{N-1} F[k] e^{j k \frac{2\pi}{N} n}, \quad n = 0, 1, \dots, N-1, \quad (1.2)$$

kde $n = 0, 1, 2, \dots, N-1$.

Rychlá Fourierova transformace (FFT - fast Fourier transform) je efektivní algoritmus pro výpočet DFT a její inverze. Výrazné zrychlení výpočtu je způsobeno použitím tzv. „motýlků“, které redukuje původní kvadratickou složitost $O(n^2)$ na logaritmicko-lineární $O(n \log_2 n)$. [40]

1.4 Artikulace

Stejně jako u mluveného slova mluvíme o artikulaci, tak je tomu obdobně i v hudbě. Artikulace nám v podstatě říká, jak na sebe jednotlivé noty navazují. Dá se říci, že popisují, co se má dít mezi notami. Opět existují značky či názvosloví pro notový zápis, které určuje způsob artikulace noty. Jsou to:



Obr. 1.4: Spektrum tónu A1 zahrané klavírem

1. **Detashe** – dá si říci, že se jedná o běžné hraní. Jen je kladen důraz na to, aby se nota zahrála v plné své délce. Není totiž výjimkou, že interpret podle svého hudebního citu nepatrně mění délku noty. V notovém zápise se značí pomlčkou nad notou.
2. **Staccato** – sekaně, tj. nota se zahraje krátce. Doba trvání noty se zkrátí přibližně na polovinu své původní hodnoty. Vnikne tím tak mezi notami mezera, avšak toto zkrácení noty nesmí nijak změnit tempo nebo rytmus skladby. Staccato se v notovém zápise značí tečkou nad notou.
3. **Legato** – vázaně ¹, tj. noty se zahrají tak, aby mezi nimi nevznikla žádná mezera. U smyčcových nástrojů se tím ve většině případů myslí to, aby se sled tónů zahrál jedním tahem smyčce. Noty, které se mají zahrát vázaně jsou v notovém zápise pod tzv. obloučkem.
4. **Marcato** – akcent, zdůraznění noty. Výsledný zvuk je obdobný jako u sforzata či sforzanda zmíněných v podkapitole dynamika. Značení v notovém zápise se používá značka > nad notou.
5. **Marcatisimo** – silný akcent, výrazné zdůraznění noty. Značení pomocí symbolu ^ nad notou.
6. **Koruna** – libovolné prodloužení noty, podle úsudku interpreta nebo dirigenta. Nejčastěji se toto označení používá na konci skladby nebo jejich částí. Značkou v notovém zápise je kombinace obloučku a tečky nad notou.

Výše uvedené artikulace lze i kombinovat.

¹V anglické literatuře se slovem legato někdy myslí detashe.



Obr. 1.5: Značení artikulace v notovém zápise

Dále existují označení nebo názvosloví artikulace pro jednotlivé nástroje (např. pro smyčcové nástroje pizzicato – noty se hrají drnkáním, nikoliv smyčcem) a nebo tzv. ozdoby jako je příraz, nátryl, trylek, vlnka, apod.[12, 32, 45]

2 ROZBOR SKLADBY

2.1 Impresionismus

Impresionismus je umělecký směr, který se objevil koncem 19. a počátku 20. století ve Francii. Nejprve se tento umělecký směr projevil ve výtvarné oblasti umění. Impresionističtí malíři jako byli např. Claude Oscar Monet, Edgar Degas, Camille Pissarro, Pierre-Auguste Renoir se ve svých obrazech snaží o co nejvěrnější zachycení nálady, dojmů, neopakovatelné události či prchavosti okamžiku. Výraz „prchavý okamžik“ je základním impresionistickým pojmem. Častými náměty obrazů se staly příroda, moře, přístavy, plachetnice, západ slunce, v případě Edgara Degase potom zachycení tanečnic a baletek. Umělci dávají přednost otevřeným prostorům, malují venku v zahradách a parcích. Důležitým a častým objektem na obraze je slunce, které prozrazuje denní dobu dění na obraze. Dávali přednost kráse přirozené před krásou ideální. Klíčovou úlohu na impresionistických obrazech hraje barva, barevnost, světelnost. To je také jedno z největších pojetí mezi výtvarným a hudebním uměním. Díla impresionistů přitáhla také pozornost i francouzských básníků jako byl Charles Baudelaire (*Květy zla*) nebo Stéphane Mallarmé.[1, 13, 42, 49]

Hudební impresionismus se tedy vyznačuje výše zmíněnou barevností, inovativními harmonickými postupy, které se již neřídí ustálenými zákony. Dále jsou pro impresionistické skladby typické časté změny modality, složité a rozmanitě obměňované akordy a uvolněná tonalita. Postrádají práci s tématem, melodie ztrácí na významu a objevuje se převážně pouze v útržkovité formě. Impresionismus v hudbě zrušil či alespoň znejasnil pravidlo formálního členění skladeb. Je to nepochybně důsledek kladení důrazu na působení okamžiku. Impresionistická skladba předpokládá u posluchače naprosté otevření se dojmu a nestojí o to, aby jí jakkoli zkoumal.[1, 16]

2.2 Claude Debussy

2.2.1 Život a dílo

Celým jménem Claude Achille Debussy (22. 8. 1862 – 25. 3. 1918) byl francouzský skladatel, který společně s Mauricem Ravelem patří k nejvýznamnějším představitelům hudebního impresionismu. Narodil se ve městě Saint-Germain-en-Laye roku 1862. Jeho otec byl obchodníkem, jeho matka švadlena. Jeho první učitelkou hry na klavír byla Marie Mauté de Fleurville – žačka Fryderyka Chopina. V roce 1872 byl přijat na pařížskou konzervatoř, kde studoval hru na klavír ve třídě profesora Marmontela.[1, 5]

Claude Debussy si vytvořil svůj vlastní postup pro kompozici. Začal ve své tvorbě experimentovat s disonancemi a intervaly, které se do té doby nepoužívaly. Jeho hudba je známa používáním netradiční tonality. Inspirací mu k tomu byla návštěva Ruska, kde po vysedávání v kavárnách se dostal do kontaktu s divokou romskou hudbou. Další inspirací mu byla mezinárodní expozice v Paříži s folkovou hudbou ze zemí jako jsou Španělsko, Čína (pentatonika, chromatika, čtvrttónová stupnice) nebo Jáva (celotónová stupnice). Všechny tyto prvky byli v evropské hudbě nevšední. Souhrn hudebních prvků a postupů, které Debussy začal používat ve své tvorbě jsou:

- používání celotónových stupnic,
- používání pentatoniky,
- používání paralelních akordů,
- bitonalita,
- opětovné použití polyfonie,
- časté používání nónových, undecimových a terdecimových akordů,
- krátké, útržkovité melodie,
- použití exotických rytmů,
- předem nepřipravené a časté modulace.[4, 30]

Velký vliv na něj měla zpěvačka Marie-Blanche Vasnier, se kterou spolupracoval. Její manžel Henri Vasnier ho seznámil s vlivnými francouzskými spisovateli té doby. Mezi ně patřil i básník Paul Verlaine, který poté napsal slova k Debussyho písním. Psal převážně kratší díla pro klavír, ale i komorní či orchestrální hudbu. Roku 1894 napsal jednu ze svých nejslavnějších skladeb, *Preludium k Faunovu odpoledni* (*Prélude à l'après-midi d'un faune*), krátkou symfonickou báseň pro orchestr, inspirovanou Mallarméovou básní. Mezi jeho díla dále patří *Bergamská suita* (*Suite bergamasque*), *Nocturnes*, *Préludes*, *Dvě arabesky* (*Deux arabesques*), *Moře* (*La Mer*) a mnoho dalších.

V roce 1909 mu diagnostikovali rakovinu tlustého střeva. Zemřel 25. března 1918 ve věku 55 let.[1, 4, 5]

2.2.2 Interpretace a technika klavírní hry

Claude Debussy interpretoval své skladby sám. Rozšířil možnosti klavírní techniky. Měl nový přístup ke klavírní zvukovosti. Debussy nechápe klavír jako kladívkový nástroj, ani jako nástroj melodický. Klavír dokáže nejlépe ze všech nástrojů umožnit harmonické pokusy a inovace, včetně míchání tónin, bitonality či polytonality (paralela k míchání barev v impresionistickém malířství).

Velmi málo pasáží má v Debussyho hudbě charakterizovat sílu nebo moc. Z ohlasů jeho interpretace víme, že byl spíše jemný nežli hlučný. Jeho projev a klavírní tón

byl popisován jako matný až zastřený, občas téměř neslyšitelný. Louis Laloy¹ tvrdil, že při interpretaci Clauda Debussyho nebylo slyšet jediný dopad kladívka na strunu, jediný záchvěv klavírní mechaniky.

Z vlastních Debussyho poznámek k interpretaci jeho děl by se hráč měl vyvarovat jakékoli romantické afektovanosti. Neměl by také za každou cenu upřednostňovat melodii. Melodie se podle skladatele v hudbě uplatní dostatečnou měrou, bez zbytečného upozorňování. Stejně tak by se hráč neměl pokoušet přílišně akcentovat některou harmonickou funkci. Velmi žádoucí je však snaha všechny složky smísit v jediný zvukový pojem. Kterýkoli aspekt, který naruší zvukovou jednotu, je pro interpretaci jeho skladeb přímo fatální. Každá melodie a každý akord má smysl jen a jedině ve své následnosti a ve spojení s ostatními zvukovými složkami díla.

Důležitý pro něj byl tzv. „klavírní dotek“. Debussy sám instruuje interprety: „Hrajte velmi citlivými konečky prstů. Akordy hrajte tak, jako byste nemohli prsty odlepit od klaviatury, jako kdyby vaše prsty byly přitahovány klávesami jako magnety“. V klavírním tónu skladeb Clauda Debussyho by neměly být žádné hrany a okraje. Specifikem Debussyho skladeb jsou také rozsáhlé pasáže v dynamice piano až pianissimo. Postup správné interpretace tichých pasáží: dopadat na klávesu nepřímou, tzn. šikmo a postupně se vnořovat do klaviatury skrze bříško prstu.

Nezpochybnitelnou součástí Debussyho hudby je samozřejmě pedál. Velmi často používá tlumící pedál. Obliba Debussyho v tomto pedálu spočívá hlavně v potlačení přílišných alikvót a následných nesourodých souzvuků.

Největší důraz kladl na následnost akordů. Hráč by se měl snažit, aby se akordy přelévaly z jednoho do druhého. Harmonie je základem impresionistické hudby. Debussy často opakuje akordické tóny v různých oktávách a vytváří tak mnohdy iluzi ozvěny nebo naopak prodloužení zvukového obrazu tónu. Často také akordy arpeggiuje². Někdy je arpeggio nepředepsané, ale přesto nutné, kvůli rozpětí akordu. Debussyho arpeggio má za úkol akord prosvětlit.[1]

2.3 Svit luny

Svit luny (*Clair de lune*) je třetí z celkem čtyř částí Bergamské suity:

1. Prélude (Moderato, F-Dur, 4/4),
2. Menuet (Andantino, A-Moll, 3/4),
3. Clair de lune (Andante, D^b-Dur, 9/8),
4. Passepied (Allegretto non troppo, F[#]-Moll, 4/4).

Svit luny je výrazně průzračná skladbička v trojdílné formě. Napodobuje vznešenost a eleganci. Při interpretaci této skladby je naprosto nezbytné dodržet techniku

¹francouzský muzikolog, spisovatel a sinolog (1874 – 1944)

²arpeggio je rozklad akordu hraný tak, že tóny nezní současně, ale postupně

klavírní hry popsanou v předcházející kapitole 2.2.2. Obzvláště důležitá je v této skladbě technika pedálu, která nesmí být příliš zahlcující a zahalující.[1]

Skladba je napsána v tónině D^b-Dur (předznamenání je šest béček) v 9/8 taktu. Tempo skladby je předepsáno *andante* (krokem, volně), čemuž přibližně odpovídá 76 až 108 BPM³. Celá skladba se pohybuje v dynamickém rozsahu **ppp** až **p**. Autor nás nabádá, abychom skladbu hráli *très expressif* tedy v překladu velmi expresivně. Zároveň nám autor říká, že se basová linka má hrát *con sordina* - s útlumem, ve kterém se uplatní tlumící pedál. V prvních čtrnácti taktech je napsána úvodní melodie, která se jednou s menší variancí opakuje. Následuje část skladby rozkládající se na dvanácti taktech, která má předepsáno, aby se hrála s uvolněným tempem – tempo rubato. Skladba je však většinou interpretována tak, že je hrána tempem rubatem v jisté míře po celou dobu jejího trvání. Tento způsob hry je v impresionismu a romantismu běžný. Po čtyřech taktech nám autor předepisuje *peu à peu cresc. et animé*, čímž interpreta nabádá, aby pozvolně zesiloval a oživoval. Tato část skladby je zakončena čtyřmi akordy s předepsaným arpeggiem.

Následuje desetitaktová část, ve které se nachází druhá melodie, která je zakončena v posledním taktu modulující z D^b-Dur do C[#]-moll. V nové tónině je předepsáno *en animant* - oživujíc, povzbuzovat. Nachází se zde vrchol skladby. V tónině C[#]-moll skladba setrvává po dobu šesti taktů a poté se vrací zpět do tóniny D^b-Dur. Spolu s návratem do původní tóniny přichází uvolnění a uklidnění skladby – *Calmato*. Po následujících osmi taktech se skladba vrací k původnímu tempu a úvodní melodii. Poté přichází závěr o délce v sedmi taktech. V těchto závěrečných taktech je předepsáno *morendo jusqu'à la fin*, což v překladu znamená „umírat“ až do konce.

Celý notový zápis se nachází příloze B. Notový zápis byl převzat z [2].

³BPM – Beats Per Minute (počet úhozů za minutu)

3 ZÍSKÁVÁNÍ HUDEBNÍ INFORMACE

Získávání hudební informace, anglicky Music Information Retrieval zkráceně pak MIR, je oblast hudebního výzkumu, která se zabývá zpracováním hudebních parametrů získaných ze zvukového signálu. Tato oblast výzkumu se začala výrazně rozvíjet v posledním desetiletí 20. století, díky zvýšení výkonu výpočetní techniky, které umožnily pokročilejší zpracování číslicových signálů. Tento technický pokrok umožnil získávání informací o rytmu, tempu, barvě zvuku, melodii či harmonii přímo z audio signálu.[31]

Oblasti využití MIR:

1. Systémy doporučení – systémy, které hledají uživatele podle podobnosti poslouchaných skladeb v historii a skladeb, které uživatel dosud neslyšel. Systém pak tyto skladby (nebo interprety) doporučí k poslechu. Tohoto využívá například streamovací služby jako Spotify nebo internetové rádio Pandora¹.
2. Vyhledávání podle zadaných kritérií – nástroj pro vyhledávání hudebního obsahu v rozsáhlých databázích.
3. Separace stop a rozpoznávání hudebních nástrojů – umožňuje například vytvoření karaoke nahrávky z nahrávky běžně dostupné.
4. Automatická hudební transkripce – převedení informací z hudební nahrávky do notového zápisu nebo MIDI souboru. Převod používá detekci výšky, počátku a délky tónu, identifikaci nástroje a detekci rytmu.
5. Automatické rozdělení do kategorií – systém pomocí informací z hudebního obsahu je schopen roztrždit hudbu například podle interpreta, žánru, nálady skladby, období vzniku nebo místo původu.
6. Generování hudebního obsahu – automatická kompozice a generování skladeb.

3.1 Parametrizace

Parametrizace nebo-li získání dat popisujících základní vlastnosti hudebního signálu lze rozdělit podle míry abstrakce na parametry:

1. nízké úrovně – např. časová obálka signálu, frekvenční spektrum, energie signálu atd.,
2. střední úrovně – např. parametry tempa a výšky tónu,
3. vysoké úrovně – např. určení tóniny, žánru, nálady. . . ,

dále pak podle časového intervalu na:

1. okamžité – určitý bod v čase (desítky milisekund),
2. úsekové – část nahrávky jako je téma, fráze nebo refrém,

¹Pandora je zatím dostupná pouze v USA

3. globální – popisují skladbu jako celek.

Parametry lze získat jak z časové oblasti (např. Amplitudovou obálku signálu) tak i z frekvenční oblasti (např. Poměr energií pásem).[20]

3.1.1 Rytmické a tempové parametry

1. Fluktuace (*fluctuation*) – je odhad rytmičké struktury skladby. Výpočet je rozdělen do tří kroků:

- Nejprve se spočítá výkonový spektrogram s oknem o velikosti 23 ms. Výsledný spektrogram je rozdělen do 20 pásem. Jedná se o kritická pásma. Jednotkou je bark. Přepočet z barků na frekvenci je dána vztahem:

$$Z_{\text{bark}}(f_{\text{kHz}}) = 13 \arctan(0,76f) + 3,5 \arctan\left(\frac{f}{7,5}\right)^2. \quad (3.1)$$

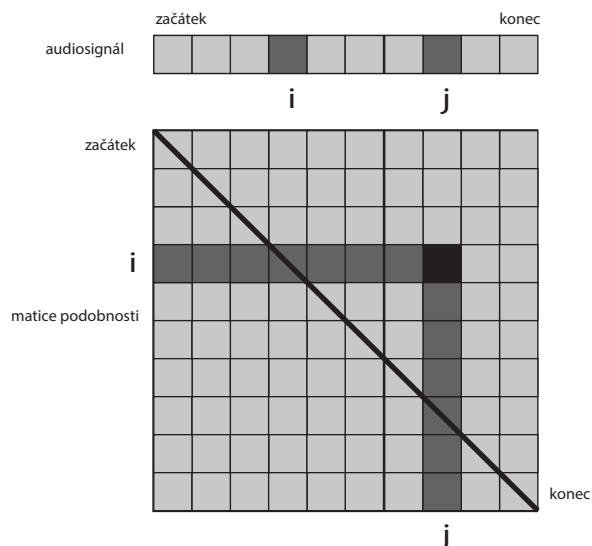
Při výpočtech je použita křivka slyšení podle Terhardta.[34]

- Pro každé pásmo je vypočteno FFT.
 - Posledním krokem je součet spekter každého pásma.[35],
2. Spektrum úderů (*beat spectrum*) – reprezentuje rytmus a tempo hudební skladby.
- Pro výpočet se nejprve vytvoří matice podobnosti **S** (similarity matrix). vytvoří se tak, že se okna (frames) signálu porovnají každé s každým. Viz obrázek 3.1. Je patrné, že největší podobnosti se nacházejí na diagonále matice, protože se zde nachází výsledek porovnání stejného okna se sebou samým.
 - Získání spektra úderů z matice podobnosti **S** pomocí autokorelace této matice podle vztahu:

$$B(k, l) = \sum_{i,j} \mathbf{S}(i, j) \mathbf{S}(i + k, j + l). \quad (3.2)$$

Jelikož výsledek autokorelace a sumace $B(k, l)$ je symetrický, lze pro získání samotného spektra úderů $B(l)$ sčítat pouze podle jedné proměnné.[10, 27]

3. Počátek notové události (*events*) – detekce maxim z amplitudové obálky. Předpokládá se, že lokální maximum, tedy nárůst energie v signálu odpovídá počátku zahrané noty nebo akordu. Pro výpočet se používá krátkodobá Furierova transformace[23],
4. Hustota notových událostí (*event density*) – průměrná četnost výskytu notových událostí. Hodnota odpovídá počtu počátků notové události za jednu sekundu[23],
5. Tempo (*tempo*) – existují dva postupy detekce tempa:



Obr. 3.1: Matice podobnosti

- Získání tempa z křivky detekce notových událostí. Odhad tempa založeném na periodicitě těchto událostí. Pro každé okno je vybrána maximální hodnota této periodicity. Tempo lze pak spočítat podle vzorce:

$$T = \frac{60B}{\delta t} [\text{BPM}], \quad (3.3)$$

kde δt je časový úsek a B je počet úderů.

- Pomocí vytvoření hierarchické metrické struktury. Ta umožňuje najít koherentní metrické úrovně, ze kterých lze pak vytvořit křivku popisující tempo skladby.

Další možností získání informace o tempu je zobrazení změn tempa (*tempo change*). Změnu tempa lze získat výpočtem rozdílu dvou po sobě jdoucích hodnot z křivky popisující tempo. Pokud je poměr hodnot tempa vyjádřen v logaritmickém měřítku o základu 2, pak zvýšení tempa odpovídá kladné hodnotě a snížení pak hodnotě záporné. Pokud se hodnota nemění výsledkem je 0.

6. Metrum (*metre*) – popisuje metrickou strukturu skladby. Vypočítá se z amplitudové obálky signálu a fluktuace spektra. Parametr je vhodný pro zobrazení rytmické složitosti skladby.[23]
7. Metrický centroid (*metrical centroid*) – popisuje metrickou aktivitu. Výpočet spočívá v porovnání centroidu s vybranými metrickými úrovněmi. Výsledná křivka metrického centroidu vyjadřuje časový vývoj metrické aktivity v hodnotách BPM. Křivku lze tedy přímo porovnávat s hodnotami BPM tempa. Vyšší hodnoty metrického centroidu vypovídají o vyšším elementárním členěním rytmiky skladby. [23],

8. Metrická síla (*metrical strength*) – parametr popisuje, zda-li je metrická hierarchie pravidelná a znatelná (čitelná) nebo naopak nepravidelná zastřená (nečitelná)[23],
9. Čitelnost pulzů (*pulse clarity*) – popisuje míru srozumitelnosti rytmičné nebo metrické pulzace celé skladby nebo jejich částí. Výpočet se provádí z křivky změny tempa v čase, která vychází z počátku notových událostí.[24],

3.1.2 Dynamické parametry

1. Efektivní hodnota signálu (*RMS*) – pro výpočet efektivní hodnoty signálu se používá tzv. plovoucí dohad s jednou časovou konstantou. Nejprve je vypočítána druhá mocnina signálu, která je pak filtrována filtrem typu dolní propust. Parametrem měřiče je časová konstanta *TAV* (Time of Average, doba průměrování).[44]
2. Špičková hodnota signálu (*peaks*) – pro výpočet špičkové hodnoty signálu se používá plovoucí odhad se dvěma časovými konstantami. Parametrem měření špičkové hodnoty jsou časové konstanty dvě: *AT* (Attack Time – doba náběhu) a *RT* (Release Time – doba poklesu).[44]
3. Amplitudová obálka (*amplitude envelope*) - znázorňuje změny špičkového signálu v závislosti na čase. Lze ji získat vícero způsoby:
 - pomocí koeficientů Fourierovy transformace v komplexím tvaru,
 - z časového průběhu reálného signálu pomocí jeho analytického signálu získaného Hilbertovou transformací,
 - výpočtem krátkodobé energie signálu,
 - nebo použitím odhadu špičkové hodnoty (tzv. sledovač obálky)[44],
4. Segmentace na pozicích tichých míst (segmentation at positions of long silences) – udává místa v signálu, kde efektivní hodnota klesne pod předem stanovenou prahovou úroveň.
5. Nízká energie (*low energy*) – využívá se k odhadu rozložení energie signálu v čase. Popisuje zda-li se s časem energie výrazně mění nebo zda-li zůstává konstantní. Jeden z možných způsobů výpočtu je z odhadu míry energie rámců, které mají menší energii, než je průměrná hodnota energie celého signálu.[50]

3.1.3 Parametry barvy zvuku

1. Frekvenční spektrum (spectrum) – získá se pomocí Fourierovy transformace. Viz kapitola 1.3.
2. Kepstrum (cepstrum) – nejčastěji se používá pro analýzu řeči. Výpočet přímé, diskrétní keprstrální transformace je realizován přibližně pomocí FFT. Pro kmi-

točty $\omega_k = k \frac{2\pi}{N}$:

$$S[k] = \ln \sum_{n=0}^{N-1} s[n] e^{-jk \frac{2\pi}{N} n}, \quad k = 0, 1, \dots, N-1. \quad (3.4)$$

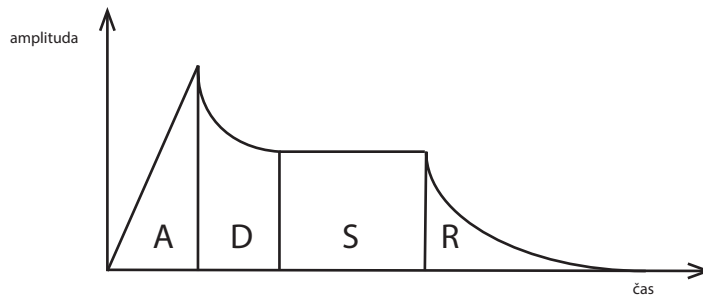
Potom je přibližný výpočet keprsta je roven:

$$s_p[n] = \frac{1}{N} \sum_{k=0}^{N-1} S[k] e^{jk \frac{2\pi}{N} n}, \quad n = 0, 1, \dots, N-1. \quad (3.5)$$

Zpětná inverzní keprstrální transformace s využitím DFT je:

$$s[n] = \frac{1}{N} \sum_{k=0}^{N-1} e^{S[k]} e^{jk \frac{2\pi}{N} n}, \quad n = 0, 1, \dots, N-1. [47] \quad (3.6)$$

3. Počet průchodů nulovou úrovní (*zero cross*) – udává, kolikrát se změní znaménku u hodnot signálu. Vychází z předpokladu, že znělá část signálu (zvuk obsahující tónovou složku) má méně průchodů nulovou úrovní nežli neznělá část signálu (šum). Používá se často při analýze řečového signálu. V případě hudebního signálu se používá pro detekci perkusivních zvuků nebo hluku. [20]
4. Doba náběhu (*attack time*) – vychází z ADSR obálky (*Attack, Decay, Sustain, Release*). Viz obrázek 3.2. Dobu náběhu lze získat z křivky počátků notových událostí. Popisuje dobu mezi tichem a maximální nabytí amplitudové velikosti zvukové vlny. [36]



Obr. 3.2: ADSR obálka

5. Doba trvání noty (*note duration*) – časový údaj, za který proběhnou všechny fáze ADSR dané noty. [36]
6. Sklon náběhu (*attack slope*) – alternativní popis první části ADSR – náběh (*attack*). Parametrem je průměrná hodnota sklonu náběhu. Hodnota je ve stejném měřítku jako původní signál, ale je normalizován v čase (čas je vyjádřen v sekundách). [36]
7. Skok náběhu (*attack leap*) – další způsob popisu náběhu. Skokem se myslí rozdíl mezi počátkem (minimální hodnota) a koncem (maximální hodnota) náběhu. Hodnota je opět vyjádřena ve stejném měřítku jako původní signál.

8. Doba útlumu (*decay time*) – obdobné jako u doby náběhu (*attack time*). Parametr však popisuje druhou část ADSR obálky.[36]
9. Sklon útlumu (*decay slope*) – analogické se sklonem náběhu, avšak pro druhou fázi ADSR.[36]
10. Skok útlumu (*decay leap*) – opět analogické se skokem náběhu, avšak pro druhou fázi ADSR.[36]
11. Pokles spektrální energie (*spectral rolloff*) – odhad míry zastoupení vysokých frekvencí v signálu. Jde o nalezení frekvence, do které je obsaženo 85% energie celkového signálu. Alternativou může být i 95% energie celkového signálu.[39, 43, 50]
12. Světlost spektra (*spectral brightness*) – popisuje procentuální zastoupení množství energie nad zvoleným kmitočtem. Nejčastěji se používá frekvence 1500 Hz. Může se však i použít 1000 Hz nebo 3000 Hz.[18]
13. Melovské keprstrální koeficienty (*Mel-Frequency Cepstrum Coefficients*) – zkratkou MFCC. Popisují zvukovou barvu. Tento parametr bere v úvahu nelineární a maskovací vlastnosti lidského slyšení. Výpočet MFCC lze rozdělit do několika bodů:

- signál je rozdělen do rámců pomocí vhodného okna (např. Hammingovo),
- rámce jsou podrobeny diskrétní Fourierovy transformaci pomocí FFT,
- převod obrazu signálu v jednotkách Hz na obraz v jednotkách mel pomocí vztahu:

$$m = 2595 \log \left(1 + \frac{f}{700} \right) = 1127 \ln \left(1 + \frac{f}{700} \right), \quad (3.7)$$

- výpočet výkonového spektra $Y[m]$ spektra signálu,
- provedení zpětné diskrétní Fourierovy transformace. Jelikož je spektrum reálná a sudá funkce, lze použít pouze diskrétní kosinovou transformaci ve tvaru:

$$c[n] = \sum_{m=0}^{N-1} Y[m] \cos \left[\frac{\pi n}{N} \left(m - \frac{1}{2} \right) \right], \quad n = 0, 1, \dots, N-1, \quad (3.8)$$

kde $C[n]$ jsou výsledné keprstrální koeficienty MFCC. [46]

14. Disonance (*roughness*) – udává míru disonance a přítomnost tzv. „rázů“ (kolísání amplitudy součtu dvou sinusovek s blízkými hodnotami jejich frekvence). Celkový odhad míry disonance v signálu se vypočítá ze špiček frekvenčního spektra. Výsledkem je průměrná hodnota disonance všech možných kombinací špiček spektra. [38, 41]
15. Nepravidelnost spektra (*irregularity*) – popisuje míru variace po sobě následujících špiček frekvenčního spektra. Existují dva způsoby výpočtu:

- podle Jensena – vzorec pro výpočet nepravidelnosti spektra J je následující:

$$J = \frac{\sum_{k=1}^N (a_k^2 - a_{k+1}^2)^2}{\sum_{k=1}^N a_k^2}, \quad (3.9)$$

kde a_k je amplituda k -té špičky.[17]

- podle Krimphoffa – vzorec pro výpočet nepravidelnosti spektra K je následující:

$$K = \sum_{k=2}^{N-1} \left| a_k - \frac{a_{k-1} + a_k + a_{k+1}}{3} \right|, \quad (3.10)$$

kde a_k je amplituda k -té špičky.[22]

16. Spektrický centroid (*spectral centroid*) – určuje střed modulového kmitočtového spektra, ve kterém je zastoupeno nejvíce energie. Používá se k měření „jasu“ zvuku (*sound brightness*). Vypočítá se podle vzorce:

$$SC_t = \frac{\sum_{n=1}^N m_t(n)n}{\sum_{n=1}^N m_t(n)}. [20] \quad (3.11)$$

17. Šířka pásma (*spectral spread, bandwidth*) – představuje pásmo spektra signálu s nejvyšší energií. Vypočítá se ze spektrálního centroidu:

$$BW_t = \frac{\sum_{n=1}^N |n - SC_t| m_t(n)}{\sum_{n=1}^N m_t(n)}. [20] \quad (3.12)$$

18. Spektrální proměnlivost (*spectral flux*) – charakterizuje vzdálenost mezi dvěma po sobě jdoucí rámce spektra. Využívá se také pro detekci řeči a počátku not. Lze vypočíst podle vztahu:

$$SF_t = \sum_{n=1}^N (D_t(n) - D_{t-1}(n))^2, \quad (3.13)$$

kde D_t je rámec po rámci normalizovaného frekvenčního rozdělení. [8, 20]

3.1.4 Statistický popis spektrálního rozložení

Parametry popisující tvar a rozložení spektra. Do této kategorie patří výše zmíněný spektrální centroid (*spectral centroid*) a šířka pásma (*spectral spread*). Dalšími parametry z této kategorie jsou:

1. Spektrální šikmost (*spectral skewness*) – popisuje symetrii spektra. Je-li výsledkem kladné číslo spektrum se naklání doleva, je-li výsledné číslo záporné naklání se spektrum doprava. Nulová hodnota vypovídá o symetrickém spektru. Často se však pro popis symetrie spektra používá koeficient šikmosti, který se spočítá pomocí vzorce:

$$C = \frac{\mu_3}{\sigma^3}, \quad (3.14)$$

kde μ_3 je spektrální šikmost a σ^3 je třetí mocnina směrodatné odchylky.[19]

2. Špičatost spektra (*spectral kurtosis*) – popisuje míru špičatosti spektra. [25]
3. Plochosť spektra (*spectral flatness*) – popisuje, zda-li je rozložení spektra hrbolaté či ploché. Spočítá se jako poměr mezi geometrického a aritmetického průměru:

$$F = \frac{\sqrt[N]{\prod_{n=0}^{N-1} x(n)}}{\frac{\sum_{n=0}^{N-1} x(n)}{N}}. [25] \quad (3.15)$$

4. Spektrální entropie (*spectral entropy*) – výsledkem je Shannonova entropie spektra. [25]

4 KLASIFIKACE

V této části práce je popsáno způsoby klasifikace a následné vyhodnocení výsledků této klasifikace.

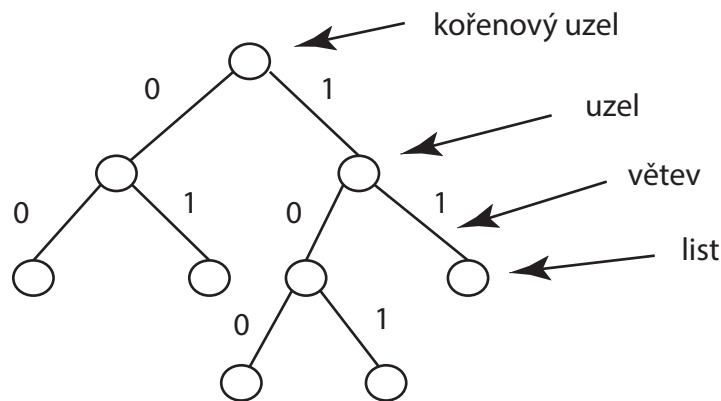
4.1 Rozhodovací strom

Rozhodovací stromové struktury (*Classification trees*) se používají v různých vědních oborech. Převážně se využívají pro klasifikaci kvalitativních závislých proměnných. V každém uzlu je dotaz na hodnotu proměnné nebo větve. V jednotlivém uzlu může být rozhodováno na základě vícerozměrné funkce tvořené kombinací vybraných nezávislých proměnných. Tato možnost bývá využívána, pokud je klasifikováno do několika málo tříd na základě mnoha nezávislých proměnných. Existují algoritmy, které umožňují automatické vytváření RS a extrahují tak znalosti z dat. Rozhodovací stromy třídí data do odlišných skupin, které tvoří nejsilnější separaci hodnot závislé proměnné. Někdy je potřeba rozhodovací strom tzv. „prořezat“, tj. nalézt optimální počet parametrů pro co nejúspěšnější klasifikaci. Výsledkem klasifikace je vyjádřen listy stromu. Nakonec se výsledky rozhodovacího stromu validují. Validace vyjadřuje předpokládanou chybu modelu.[14]

Terminologie rozhodovacích stromů:

- Atribut – vlastnost, parametr, která popisuje nebo charakterizuje sledovaný objekt.
- Uzel – je místo rozhodovacího stromu, kde při průchodu dat jsou tyto data rozdělena na základě podmínky. Dochází zde tedy k větvení stromu.
- Kořenový uzel – je počáteční uzel rozhodovacího stromu.
- List – je součást rozhodovacího stromu, kde při jeho dosažení vede ke klasifikaci nebo predikci hodnoty výstupní veličiny analyzovaných vstupních dat.
- Větev – spojnice mezi dvěma uzly nebo uzlem a listem.
- Hloubka stromu – vyjadřuje pomocí číselné hodnoty počet větví, které vedou od kořenového uzlu k nejvzdálenějšímu listu.[14]

Existuje několik různých typů rozhodovacích stromů. Liší se např. typem vstupních a výstupních dat. Pro účely klasifikace v této práci byl použit rozhodovací strom typu CART (*Classification and Regression Trees*). Stromy vytvořené touto metodou mají binární typologii. Z každého uzlu tedy vycházejí dvě větve nebo listy. Vzorky u CART se rozdělí na trénovací a testovací. Trénovací data vytvoří klasifikační strom. Testovacími daty se pak určuje přesnost (*accuracy*) s jakou daný strom dokáže správně klasifikovat data. Tedy procentuální vyjádření poměru správně a nesprávně zařazeného vzorku do kategorie. Vytváření stromu probíhá od kořene klistům.[14]



Obr. 4.1: Rozhodovací binární strom

4.2 Křížová validace

V některých případech je pro vytvoření stromu nedostatek dat. Tento nedostatek lze obejít pomocí křížové validace (*Cross-validation*). Tato metoda spočívá v tom, že pro vytvoření stromu se použijí data jak trénovací tak i testovací. Existují dvě varianty:

1. *leave-one-out* – vybere se jeden prvek, který je určen pro testování. Zbytek prvků je použit pro natrénování stromu. Tento postup se opakuje pro každý z prvků. Nevýhodou této metody je vyšší výpočetní náročnost.
2. *k-fold* – prvky se rozloží do k podmnožin. Proces je obdobný jako v předchozím případě avšak pro testování se použije k podmnožina a zbytek podmnožin slouží k trénování stromu. Postup je opakován pro každou podmnožinu k .

4.3 Minimální redundance, maximální relevance

Při klasifikaci s velkým obsahem parametrů dochází k navýšení výpočetní náročnosti a času pro jejich vyhodnocování. Ne vždy platí, že vyšší počet parametrů vede k lepším výsledkům. V některých případech totiž dochází k jevu, kdy vyšší počet parametrů nevede k úspěšnějšímu vytvoření klasifikačního stromu. Může dokonce nastat i situace, kdy vyšší počet parametrů dojde ke zhoršení klasifikace a naopak menší počet parametrů vede ke klasifikaci úspěšnější. Dochází k tzv. „přeučení“ rozhodovacího stromu. Proto je důležitým krokem klasifikace výběr parametrů s nejvyšší schopností diferencovat vstupní data, tj. vstupní data nesoucí vyšší hodnotu informace. Jedním ze způsobů k tomuto výběru parametrů je metoda mRMR – minimální redundance, maximální relevance (*minimum Redundancy Maximum Relevance*).

Cílem této metody je tedy najít sadu S parametrů x_i , které mají největší schopnost zařadit data určené ke klasifikaci do třídy c . Tedy výpočet tzv. maximální závislosti (*Max-Dependency*) podle vztahu:

$$\max D(S, c), \quad D = I(x_i, i = 1, \dots, m; c). \quad (4.1)$$

Maximální závislost je však obtížně implementovatelná. Proto se pro výpočet používá maximální relevance a minimální redundanci. [37]

Postup výpočtu je následující:

1. Nejprve se spočte maximální relevance podle vzorce:

$$\max D(S, c), \quad D = \frac{1}{|S|} \sum_{x_i \in S} I(x_i; c), \quad (4.2)$$

kde $I(x_i; c)$ je největší vzájemná výměna informací ve třídě c a S je hledaná sada funkcí.

2. Poté výpočet minimální redundanci, podle vztahu:

$$\min R(S), \quad R = \frac{1}{|S|^2} \sum_{x_i x_j \in S} I(x_i; x_j). \quad (4.3)$$

3. Následnou kombinací předešlých vztahů můžeme definovat:

$$\max \phi(D, R), \quad \phi = D - R. \quad (4.4)$$

Výsledkem je tedy kritérium *minimum Redundancy Maximum Relevance*. [6, 7, 37]

4.4 Vyhodnocování

V této části je popsány postupy vyhodnocení dat.

4.4.1 Přesnost

Přesnost ACC (*accuracy*) je jeden ze základních parametrů popisujících výsledek klasifikace. Parametr udává, kolik prvků bylo správně zařazeno do odpovídající třídy a kolik jich bylo zařazeno nesprávně. Vyjádření se udává v procentech. Přesnost lze spočítat podle vzorce:

$$\text{ACC} = \frac{A + D}{A + B + C + D} \cdot 100, \quad (4.5)$$

kde A je počet správně zařazených prvků do první třídy, B je počet nesprávně zařazených prvků do první třídy, C je počet správně zařazených prvků do druhé třídy a D je počet nesprávně zařazených prvků do druhé třídy.

4.4.2 Senzitivita

Senzitivita (SEN) vyjadřuje s jakou pravděpodobností je možné určit, zda-li byl prvek správně zařazen do první třídy. Lze vyjádřit dle vztahu:

$$\text{SEN} = \frac{A}{A + B} \cdot 100. \quad (4.6)$$

4.4.3 Specifická

Specifická (SPE) vyjadřuje s jakou pravděpodobností je možné určit, zda-li byl prvek správně zařazen do druhé třídy. Lze vyjádřit dle vztahu:

$$\text{SPE} = \frac{D}{C + D} \cdot 100. \quad (4.7)$$

4.4.4 Kompromis mezi senzitivitou a specifickou

Dalším parametrem, který popisuje výsledky klasifikace je tzv. kompromis mezi senzitivitou a specifickou (*trade-off between sensitivity and specificity*) TSS. Lze vypočítat pomocí SEN a SPE a to [48]:

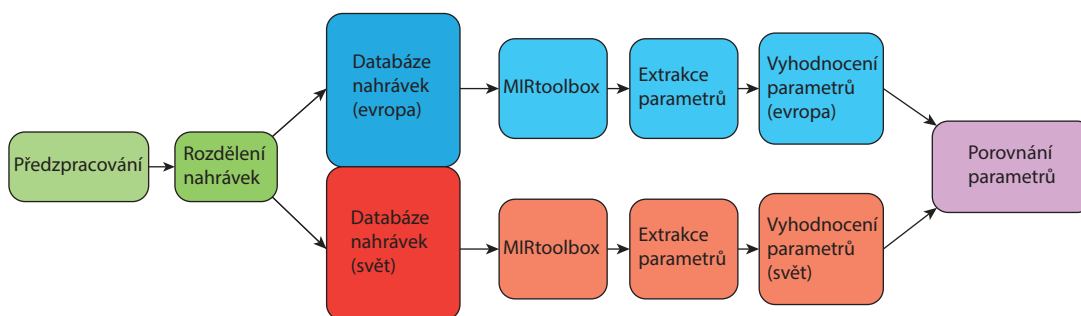
$$\text{TSS} = 2^{\sin(\frac{\pi \text{SEN}}{2}) \sin(\frac{\pi \text{SPE}}{2})}. \quad (4.8)$$

TSS hodnota leží v intervalu 1 až 2, kde hodnota blíží se hodnotě 2 vyjadřuje lepší výsledek.

5 NÁVRH ANALÝZY NAHRÁVEK

Nejčastěji se k analýze používají digitální zvukové formáty jako WAV, FLAC, ALAC, apod. Zvukové formáty se ztrátovou kompresí jako jsou např. MP3 nebo OGG jsou pro potřebnou analýzu nedostačující, zejména kvůli chybějícím informacím ve spektrální oblasti. Mohlo by to do následné analýzy zavést chyby a tím pádem získat zavádějící informace. Avšak existují výzkumy používající databanku s formátem MP3, díky jeho oblíbenosti a širokému rozšíření.

Na obrázku 5.1 je zobrazeno blokové schéma analýzy.



Obr. 5.1: Návrh analýzy nahrávek

5.1 Databáze

Databáze je rozdělena na dvě části. V první části se nachází nahrávky, které mají původ vzniku v Evropě. V druhé pak mají původ vzniku ve zbytku světa. Ruští interpreti jsou započítáni do druhé kategorie. Celkem je v databázi 59 nahrávek, z toho 38 pochází z Evropy a 21 ze zbytku světa. Přehled interpretů a jejich národností jsou v příloze v tabulce A.1.

5.2 Předzpracování

Všechny získané zvukové nahrávky byly ve formátu FLAC. Jelikož je pro provedení analýzy pomocí MIRtoolboxu potřeba formát souboru WAV, je nejprve nutné soubory na tento formát převést. Dále je potřeba zajistit, aby všechny zvukové soubory byly zaznamenány ve stejné vzorkovací frekvenci a to $f_{vz} = 44100$ Hz. Pro zpracování je také potřeba, aby vstupní nahrávky byly v mono záznamu. Mono nahrávku lze získat sumací kanálů stereo signálu.

5.3 MIRtoolbox a extrakce parametrů

MIRtoolbox je nástroj v prostředí Matlab určený k získávání hudebních informací z audio souborů. Byl vyvinut Olivierem Lartillotem a Petrim Toivianinem z finské University of Jyväskylä. Obsahuje funkce pro získání parametrů popisující výšky tónu, rytmu i barvy zvuku. Zahrnuje také banky filtrů, MFCC i funkce pro statistickou analýzu. MIRtoolbox je uživatelsky přívětivý nástroj, který umožňuje snadno získat jak nízkoúrovňové tak i parametry vysokoúrovňové. Stejná syntaxe může být použita pro analýzu jak jednotlivých zvukových souborů, tak i řady audio segmentů nebo vícekanálových signálů. Data a metody jsou uspořádány v objektově orientované architektuře.[26, 28]

Pomocí tohoto nástroje byly extrahovány parametry, které byly popsány v podkapitole 3.1. U některých parametrů jako např. spektrum, cepstrum, roughness atd. nabízí MIRtoolbox možnost „frame“, který způsobí rozdělení parametrizovaného signálu na rámce pomocí Hammingovým oknem s velikostí 50 ms a překrytím 50%. Pro tento typ parametrů byl výpočet proveden pro signál jako celek a poté i s možností rozdělení signálu na rámce. U parametrů, které byly vyjádřeny vektorem, byly navíc vypočteny tyto parametry:

1. střední hodnota (*mean*),
2. medián (*median*),
3. maximální hodnota (*max*),
4. minimální hodnota (*min*),
5. směrodatná odchylka (*std – standard deviation*),
6. 1. percentil (*1st percentile*),
7. 99. percentil (*99th percentile*),
8. mezipercentilový rozsah (*interpercentile range*),
9. 1. kvartil (*1st quartile*),
10. 3. kvartil (*3rd quartile*).

Celkem bylo takto získáno 231 parametrů.

Následně bylo pomocí metody mRMR vybráno 100 parametrů s největší schopností diferencovat sledované nahrávky.

5.4 Vyhodnocování a porovnání

Pro vytvoření rozhodovacího stromu byl použit software pro sekvenční selekci příznaku – SFFS vytvořený v Signal Processing laboratory. Software SFFS disponuje možností výběru parametrů na základě klasifikace, kde poskytuje 6 různých klasifikačních technik a také na základě regrese, kde je podporován i algoritmus klasifikačních a regresních stromů. Software obsahuje také vícero metrik pro hodnocení kvality

výběru: 18 metrik pro klasifikaci a 10 metrik pro regresi. Software navíc obsahuje i funkce potřebné pro *cross-validaci* (*k-fold*, *leave-one-out*) při procesu dopředné selekce.[11, 29]

Datový vstup pro software SFFS bylo 100 parametrů vybranými metodou mRMR. Volitelné vstupní parametry výpočtu SFFS byly nastaveny na hodnoty:

- volba počtu křížových validací (*num_cross_val*: 10)
- volba počtu k podmnožin (*k-fold*: 10)
- volba klasifikátoru – rozhodovací strom (*classifier*: *tree*)
- volba nastavení maximálního počtu parametrů (*N*: 100)
- výpočet TSS – (score: *tss*)

Výsledek výpočtu SFFS jsou hodnoty v tabulce 5.1. ¹ Kombinací těchto osmi para-

Tab. 5.1: Tabulka výsledků

Parametr	ACC [%]	SEN [%]	SPE [%]	TSS [-]
mirflatness(x)	69,83	63,81	73,16	1,38
mircentroid(x)	70,85	62,38	75,53	1,39
1p mirrolloff(x,frame,threshold)	72,20	71,43	72,63	1,43
max mirbrightness(x,frame)	75,42	76,19	75,00	1,49
min mirtempo(x,frame)	77,29	76,67	77,63	1,51
3q mirzerocross(x,frame)	78,64	74,76	80,79	1,52
1p mirrolloff(x,frame)	78,98	75,24	81,05	1,53
mirspread(x)	79,32	77,62	80,26	1,54

metrů uvedených v tabulce se dosáhlo nejlepší úrovně klasifikace.

Tři z těchto osmi parametrů popisují tu část spektra, kde se nacházejí vyšší frekvence. Jsou to parametry: pokles spektrální energie (*rolloff*) a to obě varianty (s hranicí zastoupení energie jak 85 % tak i 95 %) a světlost spektra (*spectral brightness*). Výsledné hodnoty parametrů jsou následující: pokles spektrální energie (85 %) – ACC = 78,98 %, SEN = 75,24 %, SPE = 81,05 %, TSS = 1,53. Pokles spektrální energie (95 %) – ACC = 72,20 %, SEN = 71,43 %, SPE = 72,63 %, TSS = 1,43. Porovnáme-li hodnoty těchto dvou parametrů jde vidět, že je na tom lépe verze s hranicí zastoupení energie 85 %. Což potvrzuje, že významnější informace se nachází v místech spektra s vyššími hodnotami frekvencí. U spektrální světlosti jsou výsledné hodnoty ACC = 75,42 %, SEN = 76,19 %, SPE = 75,00 % a TSS = 1,49. Výběr těchto parametrů mohlo zapříčinit odlišný způsob artikulace interpretů. Což u klavíru poukazuje na způsob akcentace jednotlivých not. Platí, že čím je nota (akord) zahrána s výraznějším akcentem, tím jsou vyšší harmonické čitelnější a je

¹1p – 1. percentil, 3q – 3. kvartil, x – zkoumaná nahrávka, threshold – hranice energetického rozsahu nastaven na 95 %

jejich zastoupení ve spektru větší. Naopak při menším akcentu se objeví výrazně utlumeně nebo se neobjeví vůbec. Výsledek mohl ovlivnit i způsob využití klavírního pedálu. Rozdíly v artikulaci interpretů jsou patrné např. hned v prvním akordu úvodní melodie a dále pak v následujících vrcholech stoupající a klesající melodie. Dále však především artikulace triol v taktu 15 až 18 a dále pak takty 20, 22 a 24. Odlišnou artikulaci lze pozorovat v místě modulace i v samotné části skladby nacházející se v tónině C[#]-moll tj. takty 37 až 42.

Není překvapením, že se mezi výsledky objevil i parametr popisující tempo skladby. Tento parametr dosáhl výsledných hodnot ACC = 77,29, SEN = 76,67, SPE = 77,63 a TSS = 1,51. Charakter skladby dává interpretovi nemalý prostor pro manipulaci s tempem skladby. Kolísání tempa je přítomno v průběhu celé skladby. Dá se předpokládat, že každý interpret zvolí jinou míru kolísání tempa a tedy dosažení i odlišných minim v tempu. Změnu tempa lze pozorovat v taktech s triolovým klesáním melodie v taktech 3 až 8. Stejně tomu tak je i na konci skladby, kde se úvodní melodie opakuje tj. takty 53 až 58. V tomto případě je však kolísání tempa ještě výraznější, než je tomu na začátku skladby. Dále pak v místě přechodu do tóniny C[#]-moll, kde dochází v některých interpretacích skladby i ke skokové změně tempa tj. takty 36 a 37. A následný přechod z prostřední části skladby zpět k úvodní melodii, takty 49 až 51. Avšak největší výkyvy tempa lze pozorovat v části skladby, kde je přímo předepsáno tempo rubato tj. takty 15 až 24.

Další z vybraných parametrů je počet průchodů nulovou úrovní (*zero cross*). Tento parametr může mít souvislost jak s dynamikou, tak i s tempem. V případě tempa to je přítomnost pauz, pomlk a tedy i odlišného frázování. Přítomnost pauz a pomlk je především v úvodní melodii – prvních 14 taktů. K výraznějšímu frázování dochází v části, kde se první melodie opět vrací. Je to způsobeno kombinací změny tempa a výrazně menší dynamiky – vznikají delší pomlky. Význam na tento parametr může mít i úplný začátek a konec skladby, kdy jednotlivý interpret vyčkává odlišnou dobu před zahájením hry a naopak na konci nechává prostor pro odeznění skladby. Výsledné hodnoty tohoto parametru jsou: ACC = 78,64 %, SEN = 74,76 %, SPE = 80,79 % a TSS = 1,52.

Poslední tři nezmíněné parametry popisují celkové rozložení energie spektra a jeho tvar. Jsou to parametry spektrický centroid (*spectral centroid*) s výslednými hodnotami: ACC = 70,85 %, SEN = 62,38 %, SPE = 75,53 %, TSS = 1,39, plochost spektra (*spectral flatness*) s výslednými hodnotami: ACC = 69,83 %, SEN = 63,81 %, SPE = 73,16 %, TSS = 1,38 a šířka pásma (*spectral spread*) s výslednými hodnotami ACC = 79,32 %, SEN = 77,62 %, SPE = 80,26 %, TSS = 1,54. Tyto parametry popisující obecný tvar spektra. Může být ovlivněno různými příčinami. Ať už to bylo použitím různých typů klavírů nebo použitím různých technik při nahrávání skladby. Odlišné způsoby snímání – použití odlišných mikrofonních technik, různá

aplikace efektů, použití odlišné hudební techniky a v neposlední řadě celkový mix nahrávky.

6 ZÁVĚR

První část této práce se věnovala rozboru impresionistické skladby *Clair de lune*, která byla zkomponována francouzským skladatelem Claudem Debussym. Zabývala se historickým zasazením skladby a zařazením do uměleckého směru. Dále byly popsány skladatelovy techniky kompozice a jeho klavírní hry spolu s poznámkami k interpretaci jeho skladeb.

V druhé části byla pozornost věnována hlavním vlastnostem hudby – tempu, dynamice a barvě zvuku. Bylo zde popsáno, jakým způsobem jsou tyto hudební vlastnosti měřeny, označovány v notovém zápise, rozlišovány a děleny a jaké názvosloví se pro jejich popis používá.

Ve třetí části bylo popsáno získávání hudební informace (MIR–Music Information Retrieval). Také se zde uvádí, ve kterých oblastech a aplikacích se tyto informace využívají. Dále se zde nachází výčet parametrů, které slouží k popisu hudebních vlastností jako je právě tempo, dynamika a barva zvuku.

Ve čtvrté části byly popsány způsoby klasifikace, rozhodovací stromy, metoda mRMR, křížová validace a parametry popisující úspěšnost klasifikace.

A v poslední části byl představen MIRtoolbox a samotný návrh extrakce a analýzy parametrů, které jsou schopny rozlišit původ interpretace skladby. Nejlepších výsledků dosáhl parametr: šířka pásma (*spectral spread*), kde bylo dosaženo hodnot: ACC = 79,32 %, SEN = 77,62 %, SPE = 80,26 %, TSS = 1,54. Nejvyšší hodnoty SPE se dosáhlo u parametru 1. percentil poklesu spektrální energie (*rolloff*) a to 81,05 %. Dosažené výsledky zaostávají před výsledky očekávanými. Očekávaný pohyb hodnot u ACC byl kolem 85% a u TSS hodnota kolem 1,7. Dosažením lepších výsledků by mohlo být rozšířením databáze nahrávek a získáním tak většího množství vstupních dat ke klasifikaci. Celkově se nejlepších výsledků dosáhlo při osmi parametrech: ACC = 79,32 %, SEN = 77,62 %, SPE = 81,05 %, TSS = 1,54.

LITERATURA

- [1] BITALOVÁ, Petra. *Claude Debussy: klavírní dílo a jeho využití na uměleckých školách*. Praha, 2011. Diplomová práce. Univerzita Karlova v Praze. Vedoucí práce doc. MgA. Věra Kopecká.
- [2] Clair de Lune. In: *Https://musescore.com* [online], 2012 [cit. 2017-12-06]. Dostupné z: <https://musescore.com/classicman/clairdelune>
- [3] ClaudeDebussyEN. *Wikipedia: the free encyclopedia* [online]. San Francisco (CA): Wikimedia Foundation, 2017 [cit. 2017-11-25]. Dostupné z: https://en.wikipedia.org/wiki/Claude_Debussy
- [4] CONNOR, Patricia. *A Critical analysis of the harmonic idiom of songs of Claude Debussy and its influence on compositions of Charles Loeffler and John Alden Carpenter*. Wichita Falls, Texas, June 1941. [online] Dostupné z: https://digital.library.unt.edu/ark:/67531/metadc798050/m2/1/high_res_d/1002503132-Connor.pdf
- [5] CZAJKOWSKÁ, Lucie. *Interpretační otázky první knihy Preludií Clauda Debussyho*. Brno, 2013. Diplomová práce. Janáčkova akademie múzických umění v Brně. Vedoucí práce prof. MgA. Alena Vlasáková.
- [6] DING, C; PENG, HC. *Proc. 2nd IEEE Computational Systems Bioinformatics Conference (CSB 2003)*, pp.523-528, Stanford, CA, Aug, 2003
- [7] DING C.; PENG HC., *Minimum redundancy feature selection from microarray gene expression data*, Journal of Bioinformatics and Computational Biology, Vol. 3, No. 2, pp.185-205, 2005. [PDF]
- [8] DIXON, S. *Onset detection revisited* [online]. 9th International Conference On Digital Audio Effects, DAFx 2006, 2006, s. 133-137. [cit.25.4.2017]. Dostupné z URL: http://www.dafx.ca/proceedings/papers/p_133.pdf
- [9] FENG Y.; ZHUANG Y.; PAN Y. (2003) *Popular music retrieval by detecting mood*, ACM SIGIR Conference on Research and Development in Information Retrieval.
- [10] FOOTE, Jonathan; COOPER, Matthew; NAM, Unjung. (2002). *Audio Retrieval by Rhythmic Similarity*. FX Palo Alto Laboratory, Inc.; Department of Music Stanford University Dostupné z URL: https://www.researchgate.net/publication/220723700_Audio_Retrieval_by_Rhythmic_Similarity

- [11] GALÁŽ, Z. *Preliminary Acoustic Analysis of Noise Components in Patients In Parkinsons Disease*. In Proceedings of the 21st Conference STUDENT EEICT 2015. Brno: 2015. p. 476-480. ISBN: 978-80-214-5148- 3.
- [12] HARNUM, Jonathan. *Basic Music Theory: How To Read, Write and Understand Written Music*. Sol-Ut Press, A Music Education Business, 2001. ISBN 0-9707512-9-X.
- [13] HEGEROVÁ, Pavla. *Estetická idea impresionismu*. Brno, 2014. Bakalářská práce. Masarykova univerzita v Brně. Vedoucí práce doc. Mgr. Rostislav Niederle PhD.
- [14] HONZÍK, Petr. *Strojové učení* [online]. Brno: Vysoké učení technické v Brně Fakulta elektrotechniky a komunikačních technologií Ústav telekomunikací, 2006.
- [15] HUDSON, Richard. *Stolen Time: the History of Tempo Rubato*. Oxford: Clarendon Press, 1994. ISBN 0-19-816169-7
- [16] ImpresionismusCZE. *Wikipedia: the free encyclopedia* [online]. San Francisco (CA): Wikimedia Foundation, 2017 [cit. 2017-11-25]. Dostupné z: [https://cs.wikipedia.org/wiki/Impresionismus_\(hudba\)](https://cs.wikipedia.org/wiki/Impresionismus_(hudba))
- [17] JENSEN, K. Kristoffler. *Timbre Models of Musical Sound*. Department of Computer Science, University of Copenhagen 1999.
- [18] JUSLIN, P. N. *Cue utilization in communication of emotion in music performance: relating performance to perception* [online]. Journal of Experimental Psychology: Human Perception and Performance, svazek 26, č. 6, 2000, s. 1797-1813 [cit.25.11.2017]. Dostupné z URL: https://psyk.uu.se/digitalAssets/510/c_510552-1_1-k_juslin_emotion2000.pdf
- [19] KOCH, R. *A Tutorial in Probability and Statistics*, [online] http://web.cecs.pdx.edu/~roy/tutorial/Stat_int.htm
- [20] KNEES, P.; SCHEDL M. *Music similarity and retrieval: an introduction to audio- and web- based strategies* [online]. 1. New York, NY: Springer Berlin Heidelberg, 2016, ISBN 978-3- 66249720-3, s. 299. [cit.25.4.2019].
- [21] KUNA, Milan a BLÁHA, Miloš. *Čas a hudba: K dramaturgii časových prostředků v hudebně interpretáčním výkonu*. Praha: nakladatelství Československé akademie věd, 1982.
- [22] KRIMPHOFF, J.; MCADAMS, Stephen; WINSBERG, S. (1994). *Caracterisation du timbre des sons complexes. II. analyses acoustiques et quantification psychophysique*. Journal De Physique. 4. 625-628.

- [23] LARTILLOT, O.; CEREGHETTI, D.; ELIARD, K.; TROST, W. J.; RAPPAZ, M.-A.; GRANDJEAN, D. *Estimating tempo and metrical features by tracking the whole metrical hierarchy* [online]. 3rd International Conference on Music and Emotion, Jyväskylä, 2013 [cit.10.12.2017].
- [24] LARTILLOT, O.; EEROLA, T.; TOIVIAINEN, P.; FORNARI, J. *Multi-feature modeling of pulse clarity: Design, validation, and optimization* [online]. 3rd International Conference on Music Information Retrieval, Philadelphia, 2008, s. 521-526 [cit.13.12.2017]
- [25] LARTILLOT, O.; *MIRtoolbox 1.7 User's Manual* [online]. University of Oslo, Department of Musicology, Norway; 3rd 2017 Dostupné z: <https://www.jyu.fi/hytk/fi/laitokset/mutku/en/research/materials/mirtoolbox/manual1-7.pdf>
- [26] LARTILLOT, O.; TOIVIAINEN, P. *A Matlab Toolbox For Musical Feature Extraction From Audio* [online]. 10th International Conference on Digital Audio Effects (DAFx-07), Bordeaux, France, September 10 -15, 2007, [cit.13.12.2017] Dostupné z: <http://dafx.labri.fr/main/papers/p237.pdf>
- [27] LI, Y.; YANG, S.; LI, W.; ZHU, Y. *Music Emotion Detecting Based on Beat Spectrum* [online]. 2010 International Conference on Computational and Information Sciences, Chengdu, 2010, s. 1245-1248. [cit.13.12.2017]
- [28] MIRtoolbox. In: <https://www.mathworks.com> [online], 2012 [cit. 2017-12-06]. Dostupné z: <https://www.mathworks.com/matlabcentral/fileexchange/24583-mirtoolbox>
- [29] MEKYSKA J., GALÁŽ Z., MŽOUREK Z., SMÉKAL Z., REKTOROVÁ I., et al. (2015) *Assessing Progress of Parkinson's Disease Using Acoustic Analysis of Phonation*. International Work Conference on Bioinspired Intelligence (IWOB 2015): 115-122.
- [30] MOSS, Charles K. *Biography (excerpts) of Claude Debussy by Charles K. Moss, M.M.Ed., M.Mus.* [online], Dostupné z: <http://jkornfeld.net/debussy.pdf>
- [31] MÜLLER, Meinard. *Fundamentals of music processing*. New York, NY: Springer Berlin Heidelberg, 2015. ISBN 978-331-9219-448.
- [32] NEUŽIL, Jiří a BENKO, Matěj. *Skripta základů hudební teorie*. Praha: Konzervatoř Jaroslava Ježka, 2010. Dostupné z: http://download.kjj.cz/pub/vyuka/2010/kl/skripta_kl.pdf

- [33] PAMPALK, E.; RAUBER, A.; MERKL, D. *A large set of audio features for sound de- scription (similarity and classification)* [online]. ACM Multimedia Juan-les-Pins, France. 2002, s. 570-579 [cit.10.12.2017]. Dostupné z URL: <http://www.ofai.at/~elias.pampalk/publications/pam_mm02.pdf>
- [34] PAMPALK, Elias; DIXON, Simon; WIDMER, Gerhard. (2002). *Exploring Music Collections by Browsing Different Views* Austrian Research Institute for Artificial Intelligence (OeFAI) Freyung 6/6,A-1010 Vienna, Austria; Department of Medical Cybernetics and Arificial Intelligence University of Vienna, Austria Dostupné z URL: http://www.ofai.at/~elias.pampalk/publications/pam_ismir03.pdf
- [35] PAMPALK, Elias; RAUBER, Andreas; MERKL, Dieter. (2002). *Content-based Organization and Visualization of Music Archives*. Proceedings of the ACM International Multimedia Conference and Exhibition. 570-579. 10.1145/641007.641121. Dostupné z URL: https://www.researchgate.net/publication/221573225_Content-based_Organization_and_Visualization_of_Music_Archives
- [36] PEETERS, G. *A large set of audio features for sound description (similarity and classification)* [online]. CUIDADO project, 2003, s. 1-25 [cit.13.12.2017]. Dostupné z URL: <http://recherche.ircam.fr/anasy/peeters/ARTICLES/Peeters_2003_cuidadoaudiofeatures.pdf>
- [37] PENG, Hanchuan; LONG, Fuhui; DING Chris, *Feature selection based on mutual information: criteria of max-dependency, max-relevance, and min-redundancy*, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 27, No. 8, pp.1226-1238, 2005. [PDF] <http://citeseerx.ist.psu.edu/viewdoc/download;jsessionid=3744A628E5F8C3F5EC7E701859FEC785?doi=10.1.1.63.5765&rep=rep1&type=pdf>
- [38] PLOMP, R.; J. M. LEVELT, W. (1965). *Tonal Consonance and Critical Bandwidth*. The Journal of the Acoustical Society of America. 38. 548-60. 10.1121/1.1909741.
- [39] POHLE, T.; PAMPALK, E.; WIDMER, G. *Evaluation of Frequently Used Audio Features for Classification of Music Into Perceptual Categories* [online]. Proceedings of the Fourth International Workshop on Content-Based Multimedia Indexing (CBMI'05), Riga, Latvia, 2005, s. 1-8 [cit.13.12.2017]

- [40] ŘÍHA, Kamil. *Pokročilé techniky zpracování obrazu* [online]. První. Brno: Vysoké učení technické v Brně Fakulta elektrotechniky a komunikačních technologií Ústav telekomunikací, 2012, 2012 [cit. 2017-12-11]. ISBN 978-80-214-4894-0. Dostupné z: https://www.vutbr.cz/www_base/priloha.php?dpid=79000
- [41] SETHARES, A. William. *Tuning, Timbre, Spectrum, Scale*. 2nd. Springer-Verlag London Berlin Heidelberg, 2005. ISBN 1-85233-797-4.
- [42] SCHAPIRO, Meyer. *Impressionism: Reflections and Perceptions*. 1st, George Braziller, October 1997 . ISBN 978-0807614204.
- [43] SCHEDL, M.; GÓMEZ E.; URBANO J. *Music Information Retrieval: Recent Developments and Applications* [online]. Foundations and Trends in Information Retrieval, svazek 8, č. 2-3, 2014, s. 127-261 [cit.25.4.2019]
- [44] SCHIMMEL, Jiří. *Studiová a hudební elektronika* [online]. Brno: Vysoké učení technické v Brně Fakulta elektrotechniky a komunikačních technologií Ústav telekomunikací, 2012, 2012 [cit. 2017-12-11]. ISBN 978-80-214-4452-2. Dostupné z: https://www.vutbr.cz/www_base/priloha.php?dpid=61387
- [45] SCHMIDT-JONES, Catherine. *Understanding Basic Music Theory* [online]. Rice University, Houston, Texas 2007, 2007 [cit. 2017-12-11].
- [46] SMEKAL, Zdeněk. *Číslíkové zpracování řeči* [online]. Brno: Vysoké učení technické v Brně Fakulta elektrotechniky a komunikačních technologií Ústav telekomunikací, 2010, 2010 [cit. 2017-12-11].
- [47] SMEKAL, Zdeněk. *Číslíkové zpracování signálů* [online]. Brno: Vysoké učení technické v Brně Fakulta elektrotechniky a komunikačních technologií Ústav telekomunikací, 2012, 2012 [cit. 2017-12-11]. ISBN 978-80-214-4639-7.
- [48] SMÉKAL, Z.; MEKYSKA, J.; GALAZ, Z.; MZOUREK, Z.; aj. *Analysis of Phonation in Patients with Parkinson's Disease using Empirical Mode Decomposition*. International Symposium on Signals, Circuits and Systems, 2015: s. 1–4.
- [49] ŠALDA, František Xaver. *Impresionismus* [online]. Dostupné z: http://volneknihy.xf.cz/Frantisek_Xaver_Salda--Impresionismus.pdf
- [50] TZANETAKIS, G.; ESSL, G.; COOK, P. *Automatic Musical Genre Classification Of Audio Signals* [online]. IEEE Transactions on Speech and Audio Processing, 2001, s. 293-302 [cit.10.12.2017].

- [51] WESTFALL, P. H. *Kurtosis as Peakedness, 1905–2014. R.I.P.* [online]. The American Statistician, svazek 68, č. 3, 2014, s. 191-195 [cit.13.12.2017].

SEZNAM SYMBOLŮ, VELIČIN A ZKRATEK

ACC	Accuracy – Přesnost
BPM	Beat Per Minute – Počet úderů za minutu
DFT	Discrete Fourier Transform – Diskrétní Furierova transformace
FFT	Fast Fourier Transform – Rychlá Fourierova transformace
MIR	Music Information Retrieval – Získávání hudební informace
MFCC	Mel-frequency cepstral coefficients – Melovské keprstrální koeficienty
mRMR	minimum Redundancy Maximum Relevance – minimální redundance maximální relevance
RMS	Root Mean Square – Efektivní hodnota signálu
SEN	Sensitivity – Sensitivita
SFFS	Software pro sekvenční selekci příznaku
SPE	Specificity – Specificita

SEZNAM PŘÍLOH

A	Tabulka klavírních interpretů	48
B	Notový zápis	49
C	Tabulka parametrů	55
D	Obsah přiloženého CD	57

A TABULKA KLAVÍRNÍCH INTERPRETŮ

Tab. A.1: Tabulka interpretů

Evropa	Svět
Aldo Ciccolini (Itálie, Francie)	Alicia Zizzo (USA)
Alexander Romanovsky (Ukrajina)	Antol Ugorski (Rusko)
Antonín Kubálek (Česko)	Angela Hewitt (Kanada)
Cristina Ariagno (Itálie)	Carol Rosenberger (USA)
Francois-Joel Thiollier (Francie)	Jane Coop (Kanada)
Neznámý interpret (Francie)	Jenny Lin (Tchaj-wan)
Gergely Boganyi (Maďarsko)	John Chen (Malajsie, Nový Zéland)
Harold Bauer (UK)	Junko Kobayashi (Japonsko)
Harriet Cohen (UK)	Konstantin Scherbakov (Rusko)
Ingolf Wunder (Rakousko)	Kun-Woo Paik (Jižní Korea)
Ivan Moravec (Česko)	Kyoko Tabe (Japonsko)
Iwona Karasinska-Schlair (Rakousko)	Maureen Volk (Kanada)
Javier Perianes (Španělsko)	Myung-whun Chung (Jižní Korea)
Jean-Efflam Bavouzet (Francie)	Natacha Kudritskaya (Rusko)
José Iturbi Báguena (Španělsko)	Nelson Freire (Brazílie)
Julius Drake (UK)	Nino Gvetadze (Gruzie)
Kasparas Uinskas (Litva)	Percy Grainger (Austrálie)
Leif Ove Andsnes (Norsko)	Roberta Rust (USA)
Livia Rév (Maďarsko)	Sally Whitwell (Austrálie)
Neznámý interpret (UK)	Takahiro Yoshikawa (Japonsko)
Martin Souter (UK)	Yakov Flier (Rusko)
Martino Tirimo (Kypr)	
Monique Haas (Francie)	
Moura Lympany (UK)	
Nicholas Ross (Irsko)	
Ottavia Maria Maceratini (Itálie)	
Pascal Roge 1 (Francie)	
Pascal Roge 2 (Francie)	
Patrik Jablonski (Švédsko)	
Peter Rösel (Německo)	
Pietro Spada (Itálie)	
Roland Peter Pöntinen (Švédsko)	
Samson Pascal François (Francie)	
Simon Trpčeski (Makedonie)	
Stephen Houghs (UK)	
Tamás Vásáry (Maďarsko)	
Tristan Pfaff (Francie)	
Walter Wilhelm Giesecking (Německo)	

B NOTOVÝ ZÁPIS

Clair de Lune

from "Suite Bergamasque" L. 75

Claude Debussy
(1862 – 1918)

3rd Movement

Andante très expressif

Piano

pp *con sordina*

Tempo rubato

pp *2* *m.g.*

18 *peu à peu cresc. et animé*

21

24 *dim. molto*

27 *Un poco mosso*

29

31

p

Measures 31 and 32 of a piano piece. The key signature has four flats (B-flat, E-flat, A-flat, D-flat). The music features a melody in the right hand with a long slur and a crescendo hairpin, and a bass line with eighth-note patterns. The dynamic is marked *p* (piano).

33

Measures 33 and 34. The melody continues with a slur and a crescendo hairpin. The bass line has a brief melodic phrase in measure 34. The dynamic remains *p*.

35

cresc.

Measures 35 and 36. The key signature changes to three flats (B-flat, E-flat, A-flat). The melody is marked with a crescendo hairpin. The dynamic is *cresc.* (crescendo).

En animant

37

più cresc.

Measures 37 and 38. The key signature changes to three sharps (F-sharp, C-sharp, G-sharp). The tempo/mood instruction "En animant" is written above the staff. The melody is marked with a crescendo hairpin. The dynamic is *più cresc.* (more crescendo).

39

Measures 39 and 40. The melody continues with a slur and a crescendo hairpin. The bass line has eighth-note patterns. The dynamic remains *più cresc.*

41

f *dim.*

43

calmato

pp

45

47

49

51 **a Tempo I**

ppp

8

53

8

2 2

55

57

59

pp

62

66 *pp* *morendo jusqu'à la fin*

68

70

C TABULKA PARAMETRŮ

Tab. C.1: Tabulka parametrů 1. část

Pořadí	Parametr	Popis
1	mirenvelope(x)	Amplitudová obálka
2	mirenvelope(x,frame)	Amplitudová obálka (rámce)
3	mirspectrum(x)	Spektrum
4	mirspectrum(x,frame)	Spektrum (rámce)
5	mirflux(s)	Rozdíly mezi rámcí spektra s
6	mircepstrum(x)	Kepstrum
7	mircepstrum(x,frame)	Kepstrum (rámce)
8	mirflux(c)	Rozdíly mezi rámcí kepstra c
9	mirlength(x)	Délka signálu v sekundách
10	mirrms(x)	RMS
11	mirrms(x,frame)	RMS (rámce)
12	mirlowenergy(x)	Nízká energie
13	mirlowenergy(x,Threshold)	Nízká energie ($t=1$)
14	mirlowenergy(x,ASR)	Nízká energie($t=5$)[9]
15	mirfluctuation(x)	Fluktuace
16	mirfluctuation(x,summary)	Fluktuace (sumace napříč rámců)
17	mirbeatspectrum(x)	Spektrum úderů
18	mirbeatspectrum(x,frame)	Spektrum úderů (rámce)
19	mirevents(x)	Detekce notových událostí
20	mireventdensity(x)	Četnost notových událostí
21	mireventdensity(x,frame)	Četnost notových událostí (rámce)
22	mirtempo(x,t)	Tempo
23	mirtempo(x,ac)	Tempo – křivka autokorelace
24	mirtempo(x,frame,t)	Tempo (rámce)
25	mirtempo(x,ac)	Tempo – křivka autokorelace (rámce)
26	mirtempo(x,Change)	Změny tempa
27	mirmetre(x)	Metrum
28	mirmetre(x,Frame)	Metrum (rámce)
29	mirmetroid(x)	Metrický centroid
30	mirpulseclarity(x)	Čitelnost pulzů
31	mirpulseclarity(x,Frame)	Čitelnost pulzů (rámce)
32	mirattacktime(events)	Doba náběhu
33	mirattackslope(events)	Sklon náběhu

Tab. C.2: Tabulka parametrů 2. část

Pořadí	Parametr	Popis
34	mirattackleap(events)	Skok náběhu
35	mirdecaytime(events)	Doba útlumu
36	mirdecayleap(events)	Skok útlumu
37	mirdecaylope(events)	Sklon útlumu
38	mirduration(events)	Doba trvání noty
39	mirzerocross(x)	Počet průchodů nulovou úrovní
40	mirzerocross(x,Frame)	Počet průchodů nulovou úrovní (rámce)
41	mirrolloff(x)	Pokles spektrální energie (85%)
42	mirrolloff(x,Threshold)	Pokles spektrální energie (95%)
43	mirrolloff(x,Frame)	Pokles spektrální energie (85%, rámce)
44	mirrolloff(x,Frame, Threshold)	Pokles spektrální energie (95%, rámce)
45	mirbrightness(x)	Světlost spektra
46	mirbrightness(x,frame)	Světlost spektra (rámce)
47	mircentroid(x)	Spektrální centroid
48	mirspread(x)	Šířka pásma
49	mirskewness(x)	Spektrální šikmost
50	mirkurtosis(x)	Špičatost spektra
51	mirflatness(x)	Plochosť spektra
52	mirentropy(x)	Spektrální entropie
53	mirmfcc(x)	MFCC
54	mirmfcc(x,frame)	MFCC (rámce)
55	mirroughness(x)	Disonance
56	mirroughness(x,frame)	Disonance (rámce)
57	mirregularity(x)	Nepravidelnost spektra (Jensen)
58	mirregularity(x,Krimphoff)	Nepravidelnost spektra (Krimphoff)
59	mirregularity(x,frame)	Nepravidelnost spektra (Jensen, rámce)
60	mirregularity(x,frame,Krimphoff)	Nepravidelnost spektra (Krimphoff, rámce)

D OBSAH PŘILOŽENÉHO CD

```
/ ..... kořenový adresář přiloženého CD
├── Matlab ..... zdrojové kódy
│   ├── code.m
│   └── codeDodatek.m
└── elDiplom.pdf
```